(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 May 2001 (10.05.2001)

PCT

(10) International Publication Number
WO 01/32875 A1

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, C12N 5/10, C07K 16/18, A61K 38/17, 31/711, 39/395, G01N 33/52, 33/53

(21) International Application Number: PCT/US00/30355

(22) International Filing Date:
3 November 2000 (03.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/163,521	4 November 1999 (04.11.1999)	US
09/705,211	2 November 2000 (02.11.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/163,521 (CIP)
Filed on	4 November 1999 (04.11.1999)
US	Not furnished (CIP)
Filed on	2 November 2000 (02.11.2000)

(71) Applicant *(for all designated States except US)*: **BIOGEN, INC.** [US/US]; 14 Cambridge Center, Cambridge, MA 02142 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CARULLI, John, P.** [US/US]; 9 Harris Drive, Southborough, MA 01772 (US). **LUKASHIN, Alexander** [RU/US]; 42 8th Street, Charlestown, MA 02129 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/32875 A1

(54) Title: TORSIN C NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: Disclosed are nucleic acids encoding torsin C (*torC*) polypeptides, as well as antibodies to TorC polypeptides and pharmaceutical compositions including same. TorC polypeptides share sequence identity to Torsin A and Torsin B polypeptides, which have been associated with peripheral neuropathies such as torsion dystonia.

TORSIN C NUCLEIC ACIDS AND POLYPEPTIDES

FIELD OF THE INVENTION

The invention generally relates to nucleic acids and polypeptides. The invention relates more particularly to nucleic acids encoding polypeptides related to torsin A and torsin B polypeptides, which are associated with peripheral neuropathies such as torsion dystonia.

BACKGROUND OF THE INVENTION

The dystonias represent a class of movement disorders characterized by contracted, twisted postures in afflicted individuals. The dystonia disorders are generally believed to arise from neuronal dysfunction, but not from neuronal degeneration *per se*.

Multiple genes have been implicated in various clinically distinct forms of dystonia. One such gene is DYT1, which encodes the torsin A polypeptide. DYT1 is associated with early onset, generalized dystonia. This dystonia form, and other genes implicated in genetic dystonia disorders, is inherited as an autosomal dominant marker which is not completely penetrant.

The DYT1 transcript is found at high levels in the substantia nigra pars compacta, hippocampus and cerebellum regions of the brain. The DYT1 message appears to increase in abundance with stress prior to death.

The torsin A polypeptide includes an amino-terminal signal sequence, a single membrane spanning region, and an ATP-binding domain. The protein is weakly homologous to a class of heat shock/chaperone proteins termed the HSP/ATPase family. Most cases of early onset dystonia are caused by a deletion of three nucleotides in the DYT1 gene. This deletion results in the loss of a glutamic acid residue in the carboxy terminus of the encoded torsin A protein.

SUMMARY OF THE INVENTION

The present invention is based, in part, upon the discovery of torsin C (*torC*), or DYT3, polynucleotide sequences and the Torsin C (TorC) polypeptides encoded by these nucleic acid sequences.

In one aspect, the invention provides an isolated nucleic acid which encodes a TorC polypeptide, or a fragment or derivative thereof. The nucleic acid can include, *e.g.*, nucleic acid sequence encoding a polypeptide at least 50% identical, or at least 90% identical, to a polypeptide comprising the amino acid sequence of Table 2 (SEQ ID NO:4). In some embodiments, the encoded polypeptide includes the amino acid sequence shown in Table 1 (SEQ ID NO:2).

In some embodiments the nucleic acid encodes a polypeptide which includes one or more amino acid sequences selected from the group consisting of GQHLA (SEQ ID NO:19), LSLHGW (SEQ ID NO:7), GTGK (SEQ ID NO:8), LHFPH (SEQ ID NO:9); FDEMDK (SEQ ID NO:10), PFLPL (SEQ ID NO:11), TFFP (SEQ ID NO:12), and GCKTV (SEQ ID NO:13).

In some embodiments, the nucleic acid sequence includes one or more of the sequences
129092 - 128848 (SEQ ID NO:14), 129505 - 129378 (SEQ ID NO:15), 130374 - 130199 (SEQ ID NO:16), 131378 -131113 (SEQ ID NO:17), and 132084 - 131934 (SEQ ID NO:18) of BAC clone having the accession No. AC007430.

In some embodiments, the nucleic acid sequence encodes a polypeptide that binds ATP.

The nucleic acid can include, e.g., a nucleic acid which includes the nucleotide sequence shown in Table 2 (SEQ ID NO:3), or the nucleotide sequence shown in Table 1 (SEQ ID NO:1).

The nucleic acid can be, *e.g.,* a genomic DNA fragment, or it can be a cDNA molecule.

Also included in the invention is a vector containing one or more of the nucleic acids described herein, and a cell containing the vectors or nucleic acids described herein.

The present invention is also directed to host cells transformed with a recombinant expression vector comprising any of the nucleic acid molecules described above.

In another aspect, the invention includes a pharmaceutical composition that includes a torC nucleic acid and a pharmaceutically acceptable carrier or diluent.

In a further aspect, the invention includes a substantially purified TorC polypeptide, *e.g.,* any of the TorC polypeptides encoded by a *torC* nucleic acid.

The invention also includes a pharmaceutical composition that includes a TorC polypeptide and a pharmaceutically acceptable carrier or diluent.

In a still further aspect, the invention provides an antibody that binds specifically to a TorC polypeptide. The antibody can be, *e.g.*, a monoclonal or polyclonal antibody. The invention also includes a pharmaceutical composition including *torC* antibody and a pharmaceutically acceptable carrier or diluent. The present invention is also directed to isolated antibodies that bind to an epitope on a polypeptide encoded by any of the nucleic acid molecules described above.

The present invention is further directed to kits comprising antibodies that bind to a polypeptide encoded by any of the nucleic acid molecules described above and a negative control antibody.

The invention further provides a method for producing a *torC* polypeptide. The method includes providing a cell containing a *torC* nucleic acid, *e.g.*, a vector that includes a *torC* nucleic acid, and culturing the cell under conditions sufficient to express the TorC polypeptide encoded by the nucleic acid. The expressed TorC polypeptide is then recovered from the cell. Preferably, the cell produces little or no endogenous TorC polypeptide. The cell can be, *e.g.*, a prokaryotic cell or eukaryotic cell.

The present invention provides a method of inducing an immune response in a mammal against a polypeptide encoded by any of the nucleic acid molecules disclosed above by administering to the mammal an amount of the polypeptide sufficient to induce the immune response.

The present invention is also directed to methods of identifying a compound that binds to TorC polypeptide by contacting the TorC polypeptide with a compound and determining whether the compound binds to the TorC polypeptide.

The present invention is also directed to methods of identifying a compound that binds a nucleic acid molecule encoding torsin C polypeptide by contacting TorC nucleic acid with a compound and determining whether the compound binds the nucleic acid molecule.

The invention further provides methods of identifying a compound that modulates the activity of a TorC polypeptide by contacting TorC polypeptide with a compound and determining whether the TorC polypeptide activity is modified.

The present invention is also directed to compounds that modulate TorC polypeptide activity identified by contacting a TorC polypeptide with the compound and determining

whether the compound modifies activity of the TorC polypeptide, binds to the TorC polypeptide, or binds to a nucleic acid molecule encoding a TorC polypeptide.

In another aspect, the invention provides a method of diagnosing a neurological condition, e.g., a peripheral neuropathy such as torsion dystonia, in a subject. The method includes providing a protein sample from the subject and measuring the amount of *torC* polypeptide in the subject sample. The amount of *torC* in the subject sample is then compared to the amount of torsin C polypeptide in a control protein sample. An alteration in the amount of torsin C polypeptide in the subject protein sample relative to the amount of torsin C polypeptide in the control protein sample indicates the subject has a neurological condition. A control sample is preferably taken from a matched individual, i.e., an individual of similar age, sex, or other general condition but who is not suspected of having a neurological condition. Alternatively, the control sample may be taken from the subject at a time when the subject is not suspected of having a neurological disorder. In some embodiments, the TorC polypeptide is detected using a *torC* antibody.

The invention is also directed to methods of inducing an immune response in a mammal against a polypeptide encoded by any of the nucleic acid molecules described above. The method includes administering to the mammal an amount of the polypeptide sufficient to induce the immune response.

In a further aspect, the invention includes a method of diagnosing a neurological condition, *e.g.*, a peripheral neuropathy such as torsion dystonia, in a subject. The method includes providing a nucleic acid sample, *e.g.*, RNA or DNA, or both, from the subject and measuring the amount of the *torC* nucleic acid in the subject nucleic acid sample. The amount of *torC* nucleic acid sample in the subject nucleic acid is then compared to the amount of *torC* nucleic acid in a control sample. An alteration in the amount of *torC* nucleic acid in the sample relative to the amount of *torC* in the control sample indicates the subject has a neurological condition.

In a further aspect, the invention includes a method of diagnosing a neurological condition in a subject. The method includes providing a nucleic acid sample from the subject and identifying at least a portion of the nucleotide sequence of a torsin C nucleic acid in the subject nucleic acid sample. The torsin C nucleotide sequence of the subject sample is then compared to a torsin C nucleotide sequence of a control sample. An alteration in the torsin C

nucleotide sequence in the sample relative to the torsin C nucleotide sequence in said control sample indicates the subject has a neurological condition.

In a still further aspect, the invention provides method of treating or preventing or delaying a neurological condition. The method includes administering to a subject in which such treatment or prevention or delay is desired a *torC* nucleic acid, a TorC polypeptide, or a *torC* antibody in an amount sufficient to treat, prevent, or delay a neurological condition in the subject.

The neurological conditions diagnosed, treated, prevented or delayed using the *torC* nucleic acid molecules, polypeptides or antibodies can be a peripheral neuropathy, *e.g.*, a torsion dystonia disorder. Other neurological disorders include, *e.g.*, hereditary essential tremor, Huntington's disease, and conditions characterized by loss of dopaminergic neurons or deficiencies in dopamine metabolism. These conditions include, *e.g.*, Parkinson's disease, or Parkinsonian syndrome, and rapid-onset-dystonia Parkinsonism.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

DETAILED DESCRIPTION OF THE INVENTION

The invention is based in part on the discovery of torC nucleic acid sequences, which encode a polypeptide sharing 44% sequence identity with torsin A and torsin B. Torsin A and Torsin B are members of a family of secreted proteins that were initially discovered because the seminal member of the family, torsin A, is mutated in patients with an inherited form of torsion dystonia (Ozelius et al., Nature Genet. 17:40-48, 1997). Torsion dystonia is a

peripheral neuropathy characterized by involuntary posturing of the trunk, neck, and occasionally the hands or feet.

A nucleic acid encoding a TorC polypeptide was identified in part by inspection of a deposited bacterial artificial chromosome (BAC) clone (accession number AC007430). The sequence of this clone, as well as fragments from this clone as recited herein, is incorporated herein in its entirety. GENESCAN analysis (Burge et al., J. Mol. Biol. 268:78-94, 1997) was used to assemble a database of predicted genes, and the predicted genes were analyzed using SIGNALP, a program deigned to identify signal peptides characteristic of secreted proteins (Nielsen et al., Protein Engineering 10:1-6, 1997).

This analysis identified a gene encoded in the "reverse" direction with respect to the orientation of the BAC clone AC007430. The identified gene contained exons at positions 129092 - 128848 (SEQ ID NO:14), 129505 - 129378 (SEQ ID NO:15), 130374 - 130199 (SEQ ID NO:16), 131378 -131113 (SEQ ID NO:17), and 132084 - 131934 (SEQ ID NO:18) in the BAC clone. A sequence based on these assembled exons was used to search the GenBank EST database, and an EST by the accession number aa502940 was identified. The clone from which this EST was derived was sequenced. The torsin C sequence in Table 1 is a composite of the sequence of this clone and sequences assembled from BAC AC007430.

Table 1 shows a human-derived DNA sequence encoding a TorC polypeptide, along with the amino acid sequence of an encoded polypeptide. The polypeptide sequence has an amino terminal signal peptide at amino acids 1-20 according to the amino acid sequence shown in Table 1 and is indicated as an underlined sequence. The amino terminal signal sequence in torC is shared with other members of the torsin gene family, and is typical of secreted proteins.

Table 1

```
atggcggctgcgacgcgcggctgccggccctggggctcgctcctcgggctgctcgggctg   60
M  A  A  A  T  R  G  C  R  P  W  G  S  L  L  G  L  L  G  L

gtctcggccgcggccgccgcctgggacctggcttccctgcgctgcaccttgggcgccttt  120
V  S  A  A  A  A  A  W  D  L  A  S  L  R  C  T  L  G  A  F
```

```
tgcgaatgcgacttccggcccgacttgccgggtctggagtgtgacctggctcagcacctg  180
C  E  C  D  F  R  P  D  L  P  G  L  E  C  D  L  A  Q  H  L

gccggccagcatctggccaaggcgctggtggtgaaggcgctgaaggcctttgtgcgggac  240
A  G  Q  H  L  A  K  A  L  V  V  K  A  L  K  A  F  V  R  D

ccagcccccaccaagccgctggtcctctccctgcacggctggaccggcaccggcaaatcc  300
P  A  P  T  K  P  L  V  L  S  L  H  G  W  T  G  T  G  K  S

tatgtcagctccctgctggcgcactacctcttccagggcggcctccgcagcccccgcgtg  360
Y  V  S  S  L  L  A  H  Y  L  F  Q  G  G  L  R  S  P  R  V

caccacttttctcccgtcctccacttcccccaccccagccacatcgagcgctacaagaag  420
H  H  F  S  P  V  L  H  F  P  H  P  S  H  I  E  R  Y  K  K

gatctgaagagctgggtccaagggaacctcactgcctgtggccgctccctcttcctcttc  480
D  L  K  S  W  V  Q  G  N  L  T  A  C  G  R  S  L  F  L  F

gatgagatggacaagatgcccccaggcctgatggaagtcctgcggcctttcctgggctcc  540
D  E  M  D  K  M  P  P  G  L  M  E  V  L  R  P  F  L  G  S

tcctgggtggtatacgggaccaattaccgcaaagccatcttcatcttcatcagcaacacg  600
S  W  V  V  Y  G  T  N  Y  R  K  A  I  F  I  F  I  S  N  T

ggtggcaagcagatcaaccaggtggcattggaggcgtggcgcagccggcgggaccgcgag  660
G  G  K  Q  I  N  Q  V  A  L  E  A  W  R  S  R  R  D  R  E

gagatcctcctgcaggagctggagccggtcatctcccgcgcggtgctggacaacccgcac  720
E  I  L  L  Q  E  L  E  P  V  I  S  R  A  V  L  D  N  P  H

catggcttctcaaactcgggcatcatggaagagcgcctcctagacgcagtggtgcccttc  780
H  G  F  S  N  S  G  I  M  E  E  R  L  L  D  A  V  V  P  F

ctcccgctccagcggcaccacgtccggcactgcgtgctcaacgagctggcccagctgggc  840
L  P  L  Q  R  H  H  V  R  H  C  V  L  N  E  L  A  Q  L  G

ctggagccaagggatgaggttgtccaggctgtgctggacagcaccaccttcttccctgaa  900
L  E  P  R  D  E  V  V  Q  A  V  L  D  S  T  T  F  F  P  E
```

```
gacgagcagctcttctcctccaacggctgcaagaccgtggcctcccgaatcgccttcttc  960
D  E  Q  L  F  S  S  N  G  C  K  T  V  A  S  R  I  A  F  F

ctctga  (SEQ ID NO:1)                                          966
L  *    (SEQ ID NO:2)
```

Table 2 shows a human-derived DNA sequence encoding a TorC polypeptide (SEQ ID NO: 3), along with the amino acid sequence of an encoded processed TorC polypeptide (SEQ ID NO:4), which lacks an amino terminal signal sequence.

Table 2

```
gtctcggccgcggccgccgcctgggacctggcttccctgcgctgcaccttgggcgccttt
V  S  A  A  A  A  A  W  D  L  A  S  L  R  C  T  L  G  A  F

tgcgaatgcgacttccggcccgacttgccgggtctggagtgtgacctggctcagcacctg
C  E  C  D  F  R  P  D  L  P  G  L  E  C  D  L  A  Q  H  L

gccggccagcatctggccaaggcgctggtggtgaaggcgctgaaggcctttgtgcgggac
A  G  Q  H  L  A  K  A  L  V  V  K  A  L  K  A  F  V  R  D

ccagccccaccaagccgctggtcctctccctgcacggctggaccggcaccggcaaatcc
P  A  P  T  K  P  L  V  L  S  L  H  G  W  T  G  T  G  K  S

tatgtcagctccctgctggcgcactacctcttccagggcggcctccgcagcccccgcgtg
Y  V  S  S  L  L  A  H  Y  L  F  Q  G  G  L  R  S  P  R  V

caccacttttctcccgtcctccacttcccccaccccagccacatcgagcgctacaagaag
H  H  F  S  P  V  L  H  F  P  H  P  S  H  I  E  R  Y  K  K

gatctgaagagctgggtccaagggaacctcactgcctgtggccgctccctcttcctcttc
D  L  K  S  W  V  Q  G  N  L  T  A  C  G  R  S  L  F  L  F

gatgagatggacaagatgcccccaggcctgatggaagtcctgcggcctttcctgggctcc
D  E  M  D  K  M  P  P  G  L  M  E  V  L  R  P  F  L  G  S
```

```
tcctgggtggtatacgggaccaattaccgcaaagccatcttcatcttcatcagcaacacg
S  W  V  V  Y  G  T  N  Y  R  K  A  I  F  I  F  I  S  N  T

ggtggcaagcagatcaaccaggtggcattggaggcgtggcgcagccggcgggaccgcgag
G  G  K  Q  I  N  Q  V  A  L  E  A  W  R  S  R  R  D  R  E

gagatcctcctgcaggagctggagccggtcatctcccgcgcggtgctggacaacccgcac
E  I  L  L  Q  E  L  E  P  V  I  S  R  A  V  L  D  N  P  H

catggcttctcaaactcgggcatcatggaagagcgcctcctagacgcagtggtgcccttc
H  G  F  S  N  S  G  I  M  E  E  R  L  L  D  A  V  V  P  F

ctcccgctccagcggcaccacgtccggcactgcgtgctcaacgagctggcccagctgggc
L  P  L  Q  R  H  H  V  R  H  C  V  L  N  E  L  A  Q  L  G

ctggagccaagggatgaggttgtccaggctgtgctggacagcaccaccttcttccctgaa
L  E  P  R  D  E  V  V  Q  A  V  L  D  S  T  T  F  F  P  E

gacgagcagctcttctcctccaacggctgcaagaccgtggcctcccgaatcgccttcttc
D  E  Q  L  F  S  S  N  G  C  K  T  V  A  S  R  I  A  F  F

ctctga  (SEQ ID NO:3)
L  *    (SEQ ID NO:4)
```

Table 3 shows the alignment of torsin C with torsin A and torsin B. There are numerous regions in the torsin C amino acid sequence that are identical to the corresponding torsin A and torsin B sequences. Examples of such sequences having four or more identical regions include, *e.g.*, GQHLA (SEQ ID NO:19), LSLHGW (SEQ ID NO:7), GTGK (SEQ ID NO:8), LHFPH (SEQ ID NO:9); FDEMDK (SEQ ID NO:10), PFLPL (SEQ ID NO:11), TFFP (SEQ ID NO:12), and GCKTV (SEQ ID NO:13).

Table 3

```
Torsin B          ------------------------------------------------------------
Torsin C          --MAAATRGCRPWG-------SLLGLLGLVSAAAAAWDLASLRCTLGAFCECDFRPDLPG
Torsin A          MKLGRAVLGLLLLAPSVVQAVEPISLGLALAGVLTGYIYPRLYCLFAECCGQKRSLSREA

Torsin B          ---DLEEKLFGQHLATEVIFKALTGFRNNKNPKKPLTLSLHGWAGTGKNFVSQIVAENLH
Torsin C          LECDLAQHLAGQHLAKALVVKALKAFVRDPAPTKPLVLSLHGWTGTGKSYVSSLLAHYLF
Torsin A          LQKDLDDNLFGQHLAKKIILNAVFGFINNPKPKKPLTLSLHGWTGTGKNFVSKIIAENIY
                     ** ::* *****. ::.:*: .* .:  *.***.******:****.:**.::*. :.
```

```
Torsin B        PKGLKSNFVHLFVSTLHFPHEQKIKLYQDQLQKWIRGNVSACANSVFIFDEMDKLHPGII
Torsin C        QGGLRSPRVHHFSPVLHFPHPSHIERYKKDLKSWVQGNLTACGRSLFLFDEMDKMPPGLM
Torsin A        EGGLNSDYVHLFVATLHFPHASNITLYKDQLQLWIRGNVSACARSIFIFDEMDKMHAGLI
                  **.*  ** * ..***** .:*  *:.:*: *::**::**..*:*:******: .*::

Torsin B        DAIKPFLDYYEQVDGVSYRKAIFIFLSNAGGDLITKTALDFWRAGRKREDIQLKDLEPVL
Torsin C        EVLRPFLGSSWVVYGTNYRKAIFIFISNTGGKQINQVALEAWRSRRDREEILLQELEPVI
Torsin A        DAIKPFLDYYDLVDGVSYQKAMFIFLSNAGAERITDVALDFWRSGKQREDIKLKDIEHAL
                :.::***.    * *..*:**:***:**:*.. *...**: **: :.**:* *:::* .:

Torsin B        SVGVFNNKHSGLWHSGLIDKNLIDYFIPFLPLEYRHVKMCVRAEMRARGSAIDEDIVTRV
Torsin C        SRAVLDNPHHGFSNSGIMEERLLDAVVPFLPLQRHHVRHCVLNELAQLGLEPRDEVVQAV
Torsin A        SVSVFNNKNSGFWHSSLIDRNLIDYFVPFLPLEYKHLKMCIRVEMQSRGYEIDEDIVSRV
                * .*::* : *: :*.::::..*:* .:*****: :*:: *:  *:   *    :::*  *

Torsin B        AEEMTFFPRDEKIYSDKGCKTVQSRLDFH--- (SEQ ID NO: 5)
Torsin C        LDSTTFFPEDEQLFSSNGCKTVASRIAFFL-- (SEQ ID NO: 2)
Torsin A        AEEMTFFPKEERVFSDKGCKTVFTKLDYYYDD (SEQ ID NO: 6)
                 :. ****.:*:::*.:***** ::: :.
```

Like other members of the torsin gene family, torsin C shares weak sequence similarity with the heat shock protein *clp*, an ATP dependent protease.

The present invention discloses torC nucleic acids, isolated nucleic acids that encode TorC polypeptide or a portion thereof, TorC polypeptides, vectors containing these nucleic acids, host cells transformed with the torC nucleic acids, anti-TorC antibodies, and pharmaceutical compositions. Also disclosed are methods of making TorC polypeptides, as well as methods of screening, diagnosing, treating conditions using these compounds, and methods of screening compounds that modulate TorC polypeptide activity.

The torsin C nucleic acids and polypeptides, as well as TorC antibodies, as well as pharmaceutical compositions discussed herein, are useful, *inter alia*, in treating neurological conditions. These neurological disorders include, *e.g.*, Parkinson's disease, Parkinsonian Syndrome, Huntington's Disease, and hereditary essential tremor. Neurological conditions can also include, *e.g.*, peripheral neuropathies such as torsion dystonia.

Sequences and Corresponding SEQ ID Numbers

The sequences and corresponding sequence identifier numbers (SEQ ID NOS)

discussed herein include the following:

SEQ ID NO	Description
SEQ ID NO:1	Human cDNA encoding TorC polypeptide having an amino terminal leader sequence
SEQ ID NO:2	Human TorC polypeptide with its amino terminal signal sequence
SEQ ID NO:3	Human cDNA encoding TorC polypeptide lacking an amino terminal signal sequence
SEQ ID NO:4	Human TorC polypeptide without its amino terminal signal sequence
SEQ ID NO:5	Human torsin B polypeptide sequence
SEQ ID NO:6	Human torsin A polypeptide sequence
	Conserved amino acid motifs between torsin A, torsin B, and torsin C
SEQ ID NO:19	GQHLA
SEQ ID NO:7	LSLHGW
SEQ ID NO:8	GTGK
SEQ ID NO:9	LHFPH
SEQ ID NO:10	FDEMDK
SEQ ID NO:11	PFLPL
SEQ ID NO:12	TFFP
SEQ ID NO:13	GCKTV
	TorC nucleic acid sequences and corresponding regions in BAC clone AC007430
SEQ ID NO:14	Nucleotides 129092 - 128848 of BAC clone AC007430
SEQ ID NO:15	Nucleotides 129505 - 129378 of BAC clone AC007430
SEQ ID NO:16	Nucleotides 130374 - 130199 of BAC clone AC007430
SEQ ID NO:17	Nucleotides 131378 -131113 of BAC clone AC007430
SEQ ID NO:18	Nucleotides 132084 - 131934 of BAC clone AC007430

torC Nucleic Acids

One aspect of the invention pertains to isolated nucleic acid molecules that encode TorC proteins or biologically active portions thereof. Also included are nucleic acid fragments sufficient for use as hybridization probes to identify *torC*-encoding nucleic acids (*e.g.*, *torC* mRNA) and fragments for use as polymerase chain reaction (PCR) primers for the amplification or mutation of *torC* nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

"Probes" refer to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt) or as many as about, *e.g.*, 6,000 nt, depending on use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. Probes may be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

An "isolated" nucleic acid molecule is one that is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated *torC* nucleic acid molecule can contain less than about 50 kb, 25 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, or a complement of any of these nucleotide sequences, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequences of SEQ ID NO:1 or SEQ ID NO:3 as a hybridization probe, *torC* nucleic acid sequences can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, eds., MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard

PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to *torC* nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having at least about 10 nt and as many as 50 nt, preferably about 15 nt to 30 nt. They may be chemically synthesized and may be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:3. In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:3, or a portion of this nucleotide sequence. A nucleic acid molecule that is complementary to the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:3 is one that is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:3 that it can hydrogen bond with little or no mismatches to the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:3, thereby forming a stable duplex.

As used herein, the term "complementary" refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term "binding" means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, Von der Waals, hydrophobic interactions, etc. A physical interaction can be either direct or indirect. Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1 or SEQ ID NO:3, *e.g.*, a fragment that can be used as a probe or primer or a fragment encoding a biologically active portion of *torC*. Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively, and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type.

Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the invention, in various embodiments, by at least about 45%, 50%, 70%, 80%, 95%, 98%, or even 99% identity (with a preferred identity of 80-99%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. See *e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below. An exemplary program is the Gap program (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison, WI) using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2: 482-489, which in incorporated herein by reference in its entirety).

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of TorC polypeptide. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the present invention, homologous nucleotide sequences include nucleotide sequences encoding for a TorC polypeptide of species other than humans, including, but not limited to, mammals, and thus can include, *e.g.*, mouse, rat, rabbit, dog, cat cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the nucleotide sequence encoding human torsin C protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in SEQ ID NO:4, as well as a polypeptide having torsin C activity. A homologous amino acid sequence does not encode the amino acid sequence of a human torsin C polypeptide.

The nucleotide sequence determined from the cloning of the human *torC* gene allows for the generation of probes and primers designed for use in identifying and/or cloning *torC* homologues in other cell types, *e.g.*, from other tissues, as well as *torC* homologues from other mammals. The probe/primer typically comprises a substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 consecutive sense strand nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3; or an anti-sense strand nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3; or of a naturally occurring mutant of SEQ ID NO:1 or SEQ ID NO:3.

Probes based on the human *torC* nucleotide sequence can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a TorC protein, such as by measuring a level of a *torC*-encoding nucleic acid in a sample of

cells from a subject *e.g.*, detecting *torC* mRNA levels or determining whether a genomic *torC* gene has been mutated or deleted.

"A polypeptide having a biologically active portion of *torC*" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically active portion of *torC*" can be prepared by isolating a portion of SEQ ID NO:1 or SEQ ID NO:3, that encodes a polypeptide having a *torC* biological activity (biological activities of the TorC proteins are described below), expressing the encoded portion of TorC protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of *torC*. For example, a nucleic acid fragment encoding a biologically active portion of *torC* can optionally include an ATP-binding domain. In another embodiment, a nucleic acid fragment encoding a biologically active portion of *torC* includes one or more regions.

***torC* variants**

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequences shown in Fig.1 or Fig. 3 due to degeneracy of the genetic code. These nucleic acids thus encode the same TorC protein as that encoded by the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:3. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:4.

In addition to the human *torC* nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:3, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of *torC* may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the *torC* gene may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a TorC protein, preferably a mammalian TorC protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the *torC* gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in *torC* that are the result of natural allelic variation and that do not alter the functional activity of *torC* are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding TorC proteins from other species, and thus that have a nucleotide sequence that differs from the human sequence of SEQ ID NO:1 or SEQ ID NO:3, are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the *torC* cDNAs of the invention can be isolated based on their homology to the human *torC* nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. For example, a soluble human *torC* cDNA can be isolated based on its homology to human membrane-bound *torC*. Likewise, a membrane-bound human *torC* cDNA can be isolated based on its homology to soluble human *torC*.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3. In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250 or 500 nucleotides in length. In another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (*i.e.,* nucleic acids encoding TorC proteins derived from species other than human) or other related sequences (*e.g.,* paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at Tm, 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those

in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (e.g., 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions is hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C. This hybridization is followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1 or SEQ ID NO:3 corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well known in the art. See, *e.g.*, Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50

mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50°C. Other conditions of low stringency that may be used are well known in the art (*e.g.*, as employed for cross-species hybridizations). See, *e.g.*, Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981, *Proc Natl Acad Sci USA 78*: 6789-6792.

Conservative mutations

In addition to naturally-occurring allelic variants of the *torC* sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, thereby leading to changes in the amino acid sequence of the encoded TorC protein, without altering the functional ability of the TorC protein. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NO:1 or SEQ ID NO:3. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of *torC* without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the TorC proteins of the present invention, are predicted to be particularly unamenable to alteration.

In addition, amino acid residues that are conserved among family members of the TorC proteins of the present invention, as indicated by the alignment presented as FIG. 2, are also predicted to be particularly unamenable to alteration. For example, TorC proteins of the present invention can contain at least one domain that is a typically conserved region in *TOR* family members, i.e., Torsin A and Torsin B proteins, and *torC* homologs. Examples of these conserved domains include, *e.g.*, GQHLA (SEQ ID NO:19), LSLHGW (SEQ ID NO:7), GTGK (SEQ ID NO:8), LHFPH (SEQ ID NO:9); FDEMDK (SEQ ID NO:10), PFLPL (SEQ ID NO:11), TFFP (SEQ ID NO:12), and GCKTV (SEQ ID NO:13). As such, these conserved domains are not likely to be amenable to mutation. Other amino acid residues, however, (*e.g.*, those that are not conserved or only semi-conserved among members of the

TorC proteins) may not be essential for activity and thus are likely to be amenable to alteration.

Another aspect of the invention pertains to nucleic acid molecules encoding TorC proteins that contain changes in amino acid residues that are not essential for activity. Such TorC proteins differ in amino acid sequence from SEQ ID NO:2 or SEQ ID NO:4, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 45% homologous to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4. Preferably, the protein encoded by the nucleic acid molecule is at least about 60% homologous to SEQ ID NO:2 or SEQ ID NO:4, more preferably at least about 70%, 80%, 90%, 95%, 98%, and most preferably at least about 99% homologous to SEQ ID NO:2 or SEQ ID NO:4.

An isolated nucleic acid molecule encoding a TorC protein homologous to the protein of SEQ ID NO:2 or SEQ ID NO:4 can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NO:1 or SEQ IDNO:3 by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in *torC* is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a *torC* coding sequence, such as by saturation mutagenesis, and the resultant mutants can

be screened for *torC* biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1 or SEQ ID NO:3, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

In one embodiment, a mutant TorC protein can be assayed for (1) the ability to form protein:protein interactions with other TorC proteins, other cell-surface proteins, or biologically active portions thereof, (2) complex formation between a mutant TorC protein and a *torC* ligand; (3) the ability of a mutant TorC protein to bind to an intracellular target protein or biologically active portion thereof; (*e.g.*, avidin proteins); (4) the ability to bind ATP; or (5) the ability to specifically bind a torsin protein antibody.

Antisense

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire *torC* coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a TorC protein of SEQ ID NO:2 or SEQ ID NO:4 or antisense nucleic acids complementary to a *torC* nucleic acid sequence of SEQ ID NO:1 or 3 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding *torC*. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the protein coding region of human *torC* corresponds to SEQ ID NO:3). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding *torC*. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding *torC* disclosed herein (*e.g.*, SEQ ID NO:1 or SEQ ID NO:3), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of *torC* mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of *torC* mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of *torC* mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a TorC protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

Ribozymes and PNA moieties

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave *torC* mRNA transcripts to thereby inhibit translation of *torC* mRNA. A ribozyme having

specificity for a *torC*-encoding nucleic acid can be designed based upon the nucleotide sequence of a *torC* DNA disclosed herein (*i.e.*, SEQ ID NO:1 or SEQ ID NO:3). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a *torC*-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, *torC* mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, *torC* gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the *torC* (*e.g.*, the *torC* promoter and/or enhancers) to form triple helical structures that prevent transcription of the *torC* gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of *torC* can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of *torC* can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of *torC* can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of *torC* can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of *torC* can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, Proc. *Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

TorC polypeptides

One aspect of the invention pertains to isolated TorC proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-TorC antibodies. In one

embodiment, native TorC proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, TorC proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a TorC protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the TorC protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of TorC protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of TorC protein having less than about 30% (by dry weight) of non-TorC protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-TorC protein, still more preferably less than about 10% of non-TorC protein, and most preferably less than about 5% non-TorC protein. When the TorC protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of TorC protein in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of TorC protein having less than about 30% (by dry weight) of chemical precursors or non-*torC* chemicals, more preferably less than about 20% chemical precursors or non-*torC* chemicals, still more preferably less than about 10% chemical precursors or non-*torC* chemicals, and most preferably less than about 5% chemical precursors or non-*torC* chemicals.

Biologically active portions of a TorC protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the TorC protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:4 that include fewer amino acids than the full length TorC proteins, and exhibit at least one activity of a

TorC protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the TorC protein. A biologically active portion of a TorC protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length.

A biologically active portion of a TorC protein of the present invention may contain at least one of the above-identified domains conserved between the torsin proteins. An alternative biologically active portion of a TorC protein may contain at least two of the above-identified domains. Another biologically active portion of a TorC protein may contain at least three of the above-identified domains. Yet another biologically active portion of a TorC protein of the present invention may contain at least four of the above-identified domains.

Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native TorC protein.

In an embodiment, the TorC protein has an amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:4 In other embodiments, the TorC protein is substantially homologous to SEQ ID NO:2 or SEQ ID NO:4 and retains the functional activity of the protein of SEQ ID NO:2 or SEQ ID NO:4, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail below. Accordingly, in another embodiment, the TorC protein is a protein that comprises an amino acid sequence at least about 45% homologous to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 and retains the functional activity of the TorC proteins of SEQ ID NO:2 or SEQ ID NO:4.

Determining homology between two or more sequences

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. See *Needleman and Wunsch* 1970 *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NO:1 or SEQ ID NO:3.

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison *(i.e.,* the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region.

Chimeric and fusion proteins

The invention also provides *torC* chimeric or fusion proteins. As used herein, a *torC* "chimeric protein" or "fusion protein" comprises a TorC polypeptide operatively linked to a non-TorC polypeptide. A "TorC polypeptide" refers to a polypeptide having an amino acid sequence corresponding to *torC*, whereas a "non-TorC polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the TorC protein, *e.g.*, a protein that is different from the TorC protein and that is derived from the same or a different organism. Within a *torC* fusion protein the TorC polypeptide can correspond to all or a portion of a TorC protein. In one embodiment, a *torC*

fusion protein comprises at least one biologically active portion of a TorC protein. In another embodiment, a *torC* fusion protein comprises at least two biologically active portions of a TorC protein. In yet another embodiment, a *torC* fusion protein comprises at least three biologically active portions of a TorC protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the TorC polypeptide and the non-TorC polypeptide are fused in-frame to each other. The non-TorC polypeptide can be fused to the N-terminus or C-terminus of the TorC polypeptide.

For example, in one embodiment a *torC* fusion protein comprises a *torC* domain operably linked to the extracellular domain of a second protein. Such fusion proteins can be further utilized in screening assays for compounds which modulate *torC* activity (such assays are described in detail below).

In yet another embodiment, the fusion protein is a GST-*torC* fusion protein in which the *torC* sequences are fused to the C-terminus of the GST (*i.e.*, glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant *torC*.

In another embodiment, the fusion protein is a TorC protein containing a heterologous signal sequence at its N-terminus. For example, the native *torC* signal sequence (*i.e.*, amino acids 1 to 20 of SEQ ID NO:2) can be removed and replaced with a signal sequence from another protein. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of *torC* can be increased through use of a heterologous signal sequence.

In yet another embodiment, the fusion protein is a *torC*-immunoglobulin fusion protein in which the *torC* sequences comprising one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The *torC*-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a *torC* ligand and a TorC protein on the surface of a cell, to thereby suppress *torC*-mediated signal transduction *in vivo*. The *torC*-immunoglobulin fusion proteins can be used to affect the bioavailability of a *torC* cognate ligand. Inhibition of the *torC* ligand/*torC* interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (*e.g.* promoting or inhibiting) cell survival. Moreover, the *torC*-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-TorC antibodies in a subject, to purify *torC* ligands, and in screening assays to identify molecules that inhibit the interaction of *torC* with a *torC* ligand.

A *torC* chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel *et al.* (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A *torC*-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the TorC protein.

torC agonists and antagonists

The present invention also pertains to variants of the TorC proteins that function as either *torC* agonists (mimetics) or as *torC* antagonists. Variants of the TorC protein can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of the TorC protein. An agonist of the TorC protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the TorC protein. An antagonist of the TorC protein can inhibit one or more of the activities of the naturally occurring form of the TorC protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the TorC protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the TorC proteins.

Variants of the TorC protein that function as either *torC* agonists (mimetics) or as *torC* antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of the TorC protein for TorC protein agonist or antagonist activity. In

one embodiment, a variegated library of *torC* variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of *torC* variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential *torC* sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of *torC* sequences therein. There are a variety of methods which can be used to produce libraries of potential *torC* variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential *torC* sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu Rev Biochem* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucl Acid Res* 11:477.

Polypeptide libraries

In addition, libraries of fragments of the TorC protein coding sequence can be used to generate a variegated population of *torC* fragments for screening and subsequent selection of variants of a TorC protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a *torC* coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the TorC protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of TorC proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors,

transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recrusive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify *torC* variants (Arkin and Yourvan (1992) PNAS 89:7811-7815; Delgrave *et al.* (1993) Protein Engineering 6:327-331).

Anti-*torC* Antibodies

An isolated TorC protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind *torC* using standard techniques for polyclonal and monoclonal antibody preparation. The full-length TorC protein can be used or, alternatively, the invention provides antigenic peptide fragments of *torC* for use as immunogens. The antigenic peptide of *torC* comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:4 and encompasses an epitope of *torC* such that an antibody raised against the peptide forms a specific immune complex with *torC*. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of *torC* that are located on the surface of the protein, *e.g.*, hydrophilic regions.

As disclosed herein, TorC protein sequence of SEQ ID NO:2 or SEQ ID NO:4, or derivatives, fragments, analogs or homologs thereof, may be utilized as immunogens in the generation of antibodies that immunospecifically-bind these protein components. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen, such as *torC*. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In a specific embodiment, antibodies to human TorC proteins are disclosed. Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies to a TorC protein sequence of SEQ ID

NO:2 or SEQ ID NO:4 or derivative, fragment, analog or homolog thereof. Some of these proteins are discussed below.

For the production of polyclonal antibodies, various suitable host animals (*e.g.*, rabbit, goat, mouse or other mammal) may be immunized by injection with the native protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly expressed TorC protein or a chemically synthesized TorC polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (*e.g.*, aluminum hydroxide), surface active substances (*e.g.*, lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as *Bacille Calmette-Guerin* and *Corynebacterium parvum*, or similar immunostimulatory agents. If desired, the antibody molecules directed against *torC* can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of *torC*. A monoclonal antibody composition thus typically displays a single binding affinity for a particular TorC protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular TorC protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see Kohler & Milstein, 1975 *Nature* 256: 495-497); the trioma technique; the human B-cell hybridoma technique (see Kozbor, *et al.*, 1983 *Immunol Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, *et al.*, 1983. *Proc Natl Acad Sci USA* 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus *in vitro* (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a TorC protein (see *e.g.,* U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see *e.g.,* Huse, *et al.,* 1989 *Science* 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a TorC protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See *e.g.,* U.S. Patent No. 5,225,539. Antibody fragments that contain the idiotypes to a TorC protein may be produced by techniques known in the art including, but not limited to: (*i*) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (*ii*) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (*iii*) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (*iv*) F_v fragments.

Additionally, recombinant anti-TorC antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Pat. No. 4,816,567; European Patent Application No. 125,023; Better *et al.*(1988) *Science* 240:1041-1043; Liu *et al.* (1987) *PNAS* 84:3439-3443; Liu *et al.* (1987) *J Immunol.* 139:3521-3526; Sun *et al.* (1987) *PNAS* 84:214-218; Nishimura *et al.* (1987) *Cancer Res* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; Shaw *et al.* (1988) *J Natl Cancer Inst* 80:1553-1559); Morrison(1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; U.S. Pat. No. 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J Immunol* 141:4053-4060.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of a TorC protein is facilitated

by generation of hybridomas that bind to the fragment of a TorC protein possessing such a domain. Antibodies that are specific for one or more domains within a TorC protein, *e.g.*, domains spanning the above-identified conserved regions of torsin family proteins, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

Anti-TorC antibodies may be used in methods known within the art relating to the localization and/or quantitation of a TorC protein (*e.g.*, for use in measuring levels of the TorC protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for TorC proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds [hereinafter "Therapeutics"].

An anti-*torC* antibody (*e.g.*, monoclonal antibody) can be used to isolate *torC* by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-*torC* antibody can facilitate the purification of natural *torC* from cells and of recombinantly produced *torC* expressed in host cells. Moreover, an anti-*torC* antibody can be used to detect TorC protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the TorC protein. Anti-TorC antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

torC Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding TorC protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include

those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, TorC proteins, mutant forms of *torC*, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of *torC* in prokaryotic or eukaryotic cells. For example, *torC* can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (1) to increase expression of recombinant protein; (2) to increase the solubility of the recombinant protein; and (3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION

TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the *torC* expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari, *et al.*, (1987) *EMBO J* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (InVitrogen Corp, San Diego, Calif.).

Alternatively, *torC* can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, SF9 cells) include the pAc series (Smith *et al.* (1983) *Mol Cell Biol* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J* 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells. See, *e.g.*, Chapters 16 and 17 of Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific

regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv Immunol* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, *e.g.*, the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to *torC* mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub *et al.*, "Antisense RNA as a molecular tool for genetic analysis," Reviews--Trends in Genetics, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either

mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, TorC protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding *torC* or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) TorC protein. Accordingly, the invention further provides methods for producing TorC protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding *torC* has been introduced) in a suitable medium such that TorC protein is produced. In another embodiment, the method further comprises isolating *torC* from the medium or the host cell.

Transgenic animals

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which *torC*-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous *torC* sequences have been introduced into their genome or homologous recombinant animals in which endogenous *torC* sequences have been altered. Such animals are useful for studying the function and/or activity of *torC* and for identifying and/or evaluating modulators of *torC* activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous *torC* gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing *torC*-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human *torC* DNA sequence of SEQ ID NO:1 or SEQ ID NO:3 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of the human *torC* gene, such as a mouse *torC* gene, can be isolated based on hybridization to the human *torC* cDNA (described further above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the *torC* transgene to direct expression of TorC protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in

U.S. Pat. Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan 1986, In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the *torC* transgene in its genome and/or expression of *torC* mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding *torC* can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a *torC* gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the *torC* gene. The *torC* gene can be a human gene (*e.g.*, SEQ ID NO:1 or SEQ ID NO:3), but more preferably, is a non-human homologue of a human *torC* gene. For example, a mouse homologue of human *torC* gene of SEQ ID NO:1 or SEQ ID NO:3 can be used to construct a homologous recombination vector suitable for altering an endogenous *torC* gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous *torC* gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous *torC* gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous TorC protein). In the homologous recombination vector, the altered portion of the *torC* gene is flanked at its 5' and 3' ends by additional nucleic acid of the *torC* gene to allow for homologous recombination to occur between the exogenous *torC* gene carried by the vector and an endogenous *torC* gene in an embryonic stem cell. The additional flanking *torC* nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector. See *e.g.*, Thomas *et al.* (1987) *Cell* 51:503 for a description of homologous recombination vectors. The vector is introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced *torC* gene has homologously recombined with the endogenous *torC* gene are selected (see *e.g.*, Li *et al.* (1992) *Cell* 69:915).

The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras. See *e.g.*, Bradley 1987, In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Curr Opin Biotechnol* 2:823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, see, *e.g.*, Lakso *et al.* (1992) *PNAS* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al.* (1997) *Nature* 385:810-813. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0 phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

Pharmaceutical Compositions

The *torC* nucleic acid molecules, TorC proteins, and anti-TorC antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a TorC protein or anti-*torC* antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier

for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be

prepared according to methods known to those skilled in the art, for example, as described in U.S. Pat. No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by any of a number of routes, *e.g.*, as described in U.S. Patent Nos. 5,703,055. Delivery can thus also include, *e.g.*, intravenous injection, local administration (see U.S. Pat. No. 5,328,470) or stereotactic injection (see *e.g.*, Chen *et al.* (1994) *PNAS* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: (a) screening assays; (b) detection assays (*e.g.*, chromosomal mapping, tissue typing, forensic biology), (c) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenomics); and (d) methods of treatment (*e.g.*, therapeutic and prophylactic). As described herein, in one embodiment, a TorC protein of the invention has the ability to bind ATP.

The isolated nucleic acid molecules of the invention can be used to express TorC protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications),

to detect *torC* mRNA (*e.g.*, in a biological sample) or a genetic lesion in a *torC* gene, and to modulate *torC* activity, as described further below. In addition, the TorC proteins can be used to screen drugs or compounds that modulate the *torC* activity or expression as well as to treat disorders characterized by insufficient or excessive production of TorC protein, e.g., peripheral neuropathies such as torsin dystonia, or production of TorC protein forms that have decreased or aberrant activity compared to *torC* wild type protein. In addition, the anti-TorC antibodies of the invention can be used to detect and isolate TorC proteins and modulate *torC* activity.

This invention further pertains to novel agents identified by the above described screening assays and uses thereof for treatments as described herein.

Screening Assays

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) that bind to TorC proteins or have a stimulatory or inhibitory effect on, for example, *torC* expression or *torC* activity.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a TorC protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc Natl Acad Sci U.S.A.* 90:6909; Erb *et al.* (1994) *Proc Natl Acad Sci U.S.A.* 91:11422; Zuckermann *et al.* (1994) *J Med Chem* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2059; Carell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2061; and Gallop *et al.* (1994) *J Med Chem* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), on chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner U.S. Pat. No. 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla *et al.* (1990) *Proc Natl Acad Sci U.S.A.* 87:6378-6382; Felici (1991) *J Mol Biol* 222:301-310; Ladner above.).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of TorC protein, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to a TorC protein determined. The cell, for example, can of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the TorC protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the TorC protein or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of TorC protein, or a biologically active portion thereof, on the cell surface with a known compound which binds *torC* to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TorC protein, wherein determining the ability of the test compound to interact with a TorC protein comprises determining the ability of the test compound to preferentially bind to *torC* or a biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of TorC protein, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the TorC protein or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of *torC*

or a biologically active portion thereof can be accomplished, for example, by determining the ability of the TorC protein to bind to or interact with a *torC* target molecule. As used herein, a "target molecule" is a molecule with which a TorC protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a TorC protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. A *torC* target molecule can be a non-*torC* molecule or a TorC protein or polypeptide of the present invention. In one embodiment, a *torC* target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.*, a signal generated by binding of a compound to a membrane-bound *torC* molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with *torC*.

Determining the ability of the TorC protein to bind to or interact with a *torC* target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the TorC protein to bind to or interact with a *torC* target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+}, diacylglycerol, IP_3, etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising a *torC*-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a TorC protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to the TorC protein or biologically active portion thereof. Binding of the test compound to the TorC protein can be determined either directly or indirectly as described above. In one embodiment, the assay comprises contacting the TorC protein or biologically active portion thereof with a known compound which binds *torC* to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TorC protein, wherein determining the ability of the test compound to interact with a TorC protein

comprises determining the ability of the test compound to preferentially bind to *torC* or biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-free assay comprising contacting TorC protein or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the TorC protein or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of *torC* can be accomplished, for example, by determining the ability of the TorC protein to bind to a *torC* target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of *torC* can be accomplished by determining the ability of the TorC protein further modulate a *torC* target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

In yet another embodiment, the cell-free assay comprises contacting the TorC protein or biologically active portion thereof with a known compound which binds *torC* to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TorC protein, wherein determining the ability of the test compound to interact with a TorC protein comprises determining the ability of the TorC protein to preferentially bind to or modulate the activity of a *torC* target molecule.

The cell-free assays of the present invention are amenable to use of both the soluble form or the membrane-bound form of *torC*. In the case of cell-free assays comprising the membrane-bound form of *torC*, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of *torC* is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)$_n$, 3-(3-cholamidopropyl)dimethylamminiol-1-propane sulfonate (CHAPS), 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl--N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either *torC* or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to

accommodate automation of the assay. Binding of a test compound to *torC*, or interaction of *torC* with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For example, GST-*torC* fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or TorC protein, and the mixture is incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of *torC* binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either *torC* or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated *torC* or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with *torC* or target molecules, but which do not interfere with binding of the TorC protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or *torC* trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the *torC* or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the *torC* or target molecule.

In another embodiment, modulators of *torC* expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of *torC* mRNA or protein in the cell is determined. The level of expression of *torC* mRNA or protein in the presence of the candidate compound is compared to the level of expression of *torC* mRNA or

protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of *torC* expression based on this comparison. For example, when expression of *torC* mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of *torC* mRNA or protein expression. Alternatively, when expression of *torC* mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of *torC* mRNA or protein expression. The level of *torC* mRNA or protein expression in the cells can be determined by methods described herein for detecting *torC* mRNA or protein.

In yet another aspect of the invention, the TorC proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, *e.g.*, U.S. Pat. No. 5,283,317; Zervos *et al.* (1993) Cell 72:223-232; Madura *et al.* (1993) J Biol Chem 268:12046-12054; Bartel *et al.* (1993) Biotechniques 14:920-924; Iwabuchi *et al.* (1993) Oncogene 8:1693-1696; and Brent WO94/10300), to identify other proteins that bind to or interact with *torC* ("*torC*-binding proteins" or "*torC*-bp") and modulate *torC* activity. Such *torC*-binding proteins are also likely to be involved in the propagation of signals by the TorC proteins as, for example, upstream or downstream elements of the *torC* pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for *torC* is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a *torC*-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with *torC*.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the *torC*, sequences, described herein, can be used to map the location of the *torC* genes, respectively, on a chromosome. The mapping of the *torC* sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, *torC* genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the *torC* sequences. Computer analysis of the *torC*, sequences can be used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual chromosomes of a given species. Only those hybrids containing the species-specific gene corresponding to the *torC* sequences will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the *torC* sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical like colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually.

The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases, will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma *et al.*, HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in McKusick, MENDELIAN INHERITANCE IN MAN, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland *et al.* (1987) *Nature*, 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the *torC* gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

Tissue Typing

The *torC* sequences of the present invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is digested

with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the present invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Pat. No. 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the *torC* sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The *torC* sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining

TorC protein and/or nucleic acid expression as well as *torC* activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant *torC* expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with TorC protein, nucleic acid expression or activity. For example, mutations in a *torC* gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with TorC protein, nucleic acid expression or activity.

Another aspect of the invention provides methods for determining TorC protein, nucleic acid expression or *torC* activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (*e.g.*, drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (*e.g.*, the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of *torC* in clinical trials.

These and other agents are described in further detail in the following sections.

Diagnostic Assays

An exemplary method for detecting the presence or absence of *torC* in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting TorC protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes TorC protein such that the presence of *torC* is detected in the biological sample. An agent for detecting *torC* mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to *torC* mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length *torC* nucleic acid, such as the nucleic acid of SEQ ID NO:1 or SEQ ID NO:3, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize

under stringent conditions to *torC* mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting TorC protein is an antibody capable of binding to TorC protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect *torC* mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of *torC* mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of TorC protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of *torC* genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of TorC protein include introducing into a subject a labeled anti-*torC* antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting TorC protein, mRNA, or genomic DNA, such that the presence of TorC protein, mRNA or genomic DNA is detected in the biological sample, and comparing the

presence of TorC protein, mRNA or genomic DNA in the control sample with the presence of TorC protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of *torC* in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting TorC protein or mRNA in a biological sample; means for determining the amount of *torC* in the sample; and means for comparing the amount of *torC* in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect TorC protein or nucleic acid.

Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant *torC* expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with TorC protein, nucleic acid expression or activity in, *e.g.*, neurological conditions such as peripheral neuropathies, *e.g.*, torsin dystonia. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant *torC* expression or activity in which a test sample is obtained from a subject and TorC protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of TorC protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant *torC* expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant *torC* expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder, such as a neurological disorder, *e.g.*, a peripheral neuropathy such as torsion dystonia. Thus, the present invention provides methods for determining whether a subject can be effectively

treated with an agent for a disorder associated with aberrant *torC* expression or activity in which a test sample is obtained and TorC protein or nucleic acid is detected (*e.g.*, wherein the presence of TorC protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant *torC* expression or activity.)

The methods of the invention can also be used to detect genetic lesions in a *torC* gene, thereby determining if a subject with the lesioned gene is at risk for, or suffers from, a neurological disorder, *e.g.*, a peripheral neuropathy such as torsion dystonia. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding a *torC*-protein, or the mis-expression of the *torC* gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of (1) a deletion of one or more nucleotides from a *torC* gene; (2) an addition of one or more nucleotides to a *torC* gene; (3) a substitution of one or more nucleotides of a *torC* gene, (4) a chromosomal rearrangement of a *torC* gene; (5) an alteration in the level of a messenger RNA transcript of a *torC* gene, (6) aberrant modification of a *torC* gene, such as of the methylation pattern of the genomic DNA, (7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a *torC* gene, (8) a non-wild type level of a *torC*-protein, (9) allelic loss of a *torC* gene, and (10) inappropriate post-translational modification of a *torC*-protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a *torC* gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *PNAS* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the *torC*-gene (see Abravaya *et al.* (1995) *Nucl Acids Res* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to a *torC* gene under conditions such that

hybridization and amplification of the *torC* gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli *et al.*, 1990, *Proc Natl Acad Sci USA* 87:1874-1878), transcriptional amplification system (Kwoh, *et al.*, 1989, *Proc Natl Acad Sci USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al*, 1988, *BioTechnology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a *torC* gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Pat. No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in *torC* can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin *et al.* (1996) *Human Mutation* 7: 244-255; Kozal *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in *torC* can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin *et al.* above. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of

parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the *torC* gene and detect mutations by comparing the sequence of the sample *torC* with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert (1977) *PNAS* 74:560 or Sanger (1977) *PNAS* 74:5463. It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve *et al.*, (1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publ. No. WO 94/16101; Cohen *et al.* (1996) *Adv Chromatogr* 36:127-162; and Griffin *et al.* (1993) *Appl Biochem Biotechnol* 38:147-159).

Other methods for detecting mutations in the *torC* gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type *torC* sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al* (1988) *Proc Natl Acad Sci USA* 85:4397; Saleeba *et al* (1992) *Methods Enzymol* 217:286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in *torC* cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T

at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a *torC* sequence, *e.g.*, a wild-type *torC* sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Pat. No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in *torC* genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl Acad Sci USA*: 86:2766, see also Cotton (1993) *Mutat Res* 285:125-144; Hayashi (1992) *Genet Anal Tech Appl* 9:73-79). Single-stranded DNA fragments of sample and control *torC* nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et*

al. (1989) *Proc Natl Acad. Sci USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al* (1992) *Mol Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc Natl Acad Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a *torC* gene.

Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which *torC* is expressed may be utilized in the prognostic assays described herein. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

Pharmacogenomics

Agents, or modulators that have a stimulatory or inhibitory effect on *torC* activity (*e.g.*, *torC* gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (*e.g.*, neurological, cancer-related or gestational disorders) associated with aberrant *torC* activity. In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship

between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of TorC protein, expression of *torC* nucleic acid, or mutation content of *torC* genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum, *Clin Exp Pharmacol Physiol*, 1996, 23:983-985 and Linder, *Clin Chem*, 1997, 43:254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug

response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of TorC protein, expression of *torC* nucleic acid, or mutation content of *torC* genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a *torC* modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring Clinical Efficacy

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of *torC* (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase *torC* gene expression, protein levels, or upregulate *torC* activity, can be monitored in clinical trials of subjects exhibiting decreased *torC* gene expression, protein levels, or downregulated *torC* activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease *torC* gene expression, protein levels, or downregulate *torC* activity, can be monitored in clinical trials of subjects exhibiting increased *torC* gene expression, protein levels, or upregulated *torC* activity. In such clinical trials, the expression or activity of *torC* and, preferably, other genes that have been implicated in, for example, a neurological disorder, can be used as a "read out" or markers of the immune responsiveness of a particular cell.

For example, genes, including *torC*, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) that modulates *torC* activity (*e.g.*, identified

in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of *torC* and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of *torC* or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, peptidomimetic, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (*i*) obtaining a pre-administration sample from a subject prior to administration of the agent; (*ii*) detecting the level of expression of a TorC protein, mRNA, or genomic DNA in the preadministration sample; (*iii*) obtaining one or more post-administration samples from the subject; (*iv*) detecting the level of expression or activity of the TorC protein, mRNA, or genomic DNA in the post-administration samples; (*v*) comparing the level of expression or activity of the TorC protein, mRNA, or genomic DNA in the pre-administration sample with the TorC protein, mRNA, or genomic DNA in the post administration sample or samples; and (*vi*) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of *torC* to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of *torC* to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant *torC* expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with

Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, (*i*) a TorC polypeptide, or analogs, derivatives, fragments or homologs thereof; (*ii*) antibodies to a *torC* peptide; (*iii*) nucleic acids encoding a *torC* peptide; (*iv*) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to a *torC* peptide) are utilized to "knockout" endogenous function of a *torC* peptide by homologous recombination (see, *e.g.*, Capecchi, 1989. *Science* 244: 1288-1292); or (*v*) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between a *torC* peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, a *torC* peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a *torC* peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant *torC* expression or activity, by administering to the subject an agent that modulates *torC* expression or at least one *torC* activity. Subjects at risk for a disease that is caused or contributed to by aberrant *torC* expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation

of symptoms characteristic of the *torC* aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of *torC* aberrancy, for example, a *torC* agonist or *torC* antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

Another aspect of the invention pertains to methods of modulating *torC* expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of TorC protein activity associated with the cell. An agent that modulates TorC protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a TorC protein, a peptide, a *torC* peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more TorC protein activity. Examples of such stimulatory agents include active TorC protein and a nucleic acid molecule encoding *torC* that has been introduced into the cell. In another embodiment, the agent inhibits one or more TorC protein activity. Examples of such inhibitory agents include antisense *torC* nucleic acid molecules and anti-TorC antibodies. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a TorC protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) *torC* expression or activity. In another embodiment, the method involves administering a TorC protein or nucleic acid molecule as therapy to compensate for reduced or aberrant *torC* expression or activity.

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

Equivalents

From the foregoing detailed description of the specific embodiments of the invention, it should be apparent that unique have been described. Although particular embodiments have been disclosed herein in detail, this has been done by way of example for purposes of illustration only, and is not intended to be limiting with respect to the scope of the appended claims which follow. In particular, it is contemplated by the inventor that various substitutions, alterations, and modifications may be made to the invention without departing from the spirit and scope of the invention as defined by the claims.

What is claimed is:

1. An isolated nucleic acid comprising a sequence encoding a polypeptide at least 50% identical to a polypeptide comprising the amino acid sequence of SEQ ID NO:4.

2. The nucleic acid of claim 1, wherein said nucleic acid encodes a polypeptide at least 90% identical to SEQ ID NO:4.

3. The nucleic acid of claim 1, wherein said nucleic acid encodes a polypeptide comprising SEQ ID NO:4.

4. The nucleic acid of claim 1, wherein said nucleic acid encodes a polypeptide comprising SEQ ID NO:2.

5. The nucleic acid of claim 1, wherein said nucleic acid encodes a polypeptide comprising at least one or more sequences selected from the group consisting of GQHLA (SEQ ID NO:19), LSLHGW (SEQ ID NO:7), GTGK (SEQ ID NO:8), LHFPH (SEQ ID NO:9); FDEMDK (SEQ ID NO:10), PFLPL (SEQ ID NO:11), TFFP (SEQ ID NO:12), and GCKTV (SEQ ID NO:13).

6. The nucleic acid of claim 1, wherein said encoded polypeptide binds ATP.

7. The nucleic acid of claim 1, wherein said nucleic acid comprises the nucleotide sequence of SEQ ID NO:3.

8. The nucleic acid of claim 1, wherein said nucleic acid comprises the nucleotide sequence of SEQ ID NO:1.

9. The nucleic acid of claim 1, wherein said nucleic acid is DNA.

10. The nucleic acid of claim 1, wherein said nucleic acid is RNA.

11. The nucleic acid of claim 1, wherein said nucleic acid is a cDNA molecule.

12. An isolated nucleic acid comprising a nucleotide sequence complementary to at least a portion of SEQ ID NO:3.

13. A complement of the nucleic acid of claim 12, wherein said nucleic acid of claim 12 encodes at least a portion of a torsin C homolog protein.

14. The nucleic acid of claim 12 wherein said molecule is an antisense oligonucleotide directed to SEQ ID NO:3.

15. A vector comprising the nucleic acid of claim 1.

16. A vector of claim 15, wherein said nucleic acid comprises SEQ ID NO:3.

17. A cell comprising the vector of claim 15.

18. A pharmaceutical composition comprising the nucleic acid of claim 1 and a pharmaceutically acceptable carrier.

19. A substantially purified polypeptide encoded by the nucleic acid of claim 1.

20. A pharmaceutical composition comprising the polypeptide of claim 19 and a pharmaceutically acceptable carrier.

21. An antibody which binds specifically to the polypeptide of claim 19.

22. An antibody that binds to an epitope on a polypeptide of claim 19.

23. A kit comprising an antibody that binds to a polypeptide of claim 19, and, optionally, a negative control antibody.

24. A pharmaceutical composition comprising the antibody of claim 21 and a pharmaceutically acceptable carrier.

25. A method of producing a torsin C polypeptide, the method comprising:

providing the cell of claim 17;

culturing said cell under conditions sufficient to express said torsin C polypeptide; and

recovering said torsin C polypeptide, thereby producing said torsin C polypeptide.

26. The method of claim 25, wherein said cell is a prokaryotic cell.

27. The method of claim 25, wherein said cell is a eukaryotic cell.

28. A method of diagnosing a neurological condition in a subject, the method comprising:

providing a protein sample from said subject;

measuring the amount of torsin C polypeptide in said subject sample; and

comparing the amount of torsin C polypeptide in said subject protein sample to the amount of torsin C polypeptide in a control protein sample, wherein an alteration in the amount of torsin C polypeptide in said subject protein sample relative to the amount of torsin C polypeptide in said control protein sample indicates the subject has a neurological condition.

29. The method of claim 28, wherein said torsin C polypeptide is detected using a TorC antibody.

30. The method of claim 28, wherein said neurological condition is a peripheral neuropathy.

31. The method of claim 28, wherein said neurological condition is torsion dystonia.

32. A method of diagnosing a neurological condition in a subject, the method comprising:

providing a nucleic acid sample from said subject;

measuring the amount of torsin C nucleic acid in said subject nucleic acid sample; and

comparing the amount of torsin C nucleic acid sample in said subject nucleic acid to the amount of torsin C nucleic acid in a control sample, wherein an alteration in the amount of torsin C nucleic acid in said sample relative to the amount of torsin C in said control sample indicates the subject has a neurological condition.

33. The method of claim 32, wherein the measured torsin C nucleic acid is torsin C RNA.

34. The method of claim 32, wherein the measured torsin C nucleic acid is torsin C DNA.

35. The method of claim 32, wherein the torsin C nucleic acid is measured using a TorC nucleic acid.

36. The method of claim 32, wherein the torsin C nucleic acid is measured by using one or more nucleic acids which amplify the nucleic acid of claim 1.

37. A method of diagnosing a neurological condition in a subject, the method comprising:

providing a nucleic acid sample from said subject;

identifying at least a portion of the nucleotide sequence of a torsin C nucleic acid in said subject nucleic acid sample; and

comparing the torsin C nucleotide sequence of said subject sample to a torsin C nucleotide sequence of a control sample, wherein an alteration in the torsin C nucleotide sequence in said sample relative to the torsin C nucleotide sequence in said control sample indicates the subject has a neurological condition.

38. A method of treating or preventing or delaying a neurological condition, the method comprising administering to a subject in which such treatment or prevention or delay is desired the nucleic acid of claim 1 in an amount sufficient to treat, prevent, or delay a neurological condition in said subject.

39. The method of claim 38, wherein said neurological condition is a peripheral neuropathy.

40. The method of claim 38, wherein said neurological condition is torsion dystonia.

41. A method of treating or preventing or delaying a neurological condition, the method comprising administering to a subject in which such treatment or prevention or delay is desired the polypeptide of claim 20 in an amount sufficient to treat, prevent, or delay a neurological condition in said subject.

42. A method of treating or preventing or delaying a neurological condition, the method comprising administering to a subject in which such treatment or prevention or delay is desired the antibody of claim 24 in an amount sufficient to treat, prevent or delay a neurological condition in said subject.

43. A method for identifying a compound that binds torsin C protein comprising the steps of:

a) contacting torsin C protein with a compound; and

b) determining whether said compound binds torsin C protein.

44. A method of claim 43, wherein binding of said compound to torsin C protein is determined by a protein binding assay.

45. A compound identified by the method of claim 43.

46. A method for identifying a compound that binds a nucleic acid encoding torsin C protein comprising the steps of:

a) contacting said nucleic acid encoding torsin C protein with a compound; and

b) determining whether said compound binds said nucleic acid molecule.

47. A compound identified by the method of claim 46.

48. A method for identifying a compound that modulates the activity of torsin C protein comprising the steps of:

a) contacting torsin C protein with a compound; of

b) determining whether torsin C protein activity has been modulated.

49. A compound identified by the method of claim 48.